白花刺续断野生居群的叶绿体全基因组特征解析

张倩[1]，张德全[1,2]*

（1. 大理大学 药学院，云南 大理 671000；2. 云南省滇西抗病原植物资源筛选研究重点实验室，云南 大理 671000）

摘要：白花刺续断是藏区一种常用的药用植物，但其叶绿体全基因组的研究较少。为揭示该物种叶绿体全基因组的基本特征并探讨其谱系遗传结构，该研究利用 Illumina 测序平台对来自 5 个野生居群的 10 个白花刺续断个体进行二代测序，经组装、注释，得到 10 条完整的叶绿体全基因组序列，并对它们的基因组特征和居群间的谱系进化关系进行了初步研究。研究结果表明：（1）白花刺续断的叶绿体全基因组大小为 155 335~156 266 bp，共注释 113 个基因，包括 72 个蛋白编码基因、30 个 tRNA 基因和 4 个 rRNA 基因，其叶绿体基因组的大小、结构、GC 含量及基因组成等方面在种内高度保守；（2）基因组比较分析表明，白花刺续断变异较大的片段均位于单拷贝区，且 IR 边界未出现明显的扩张和收缩；（3）群体遗传分析发现，白花刺续断的野生居群具有明显的地理遗传结构，不同居群间在遗传距离与地理距离上具有一定的相关性。该研究表明，白花刺续断叶绿体基因组在种内居群水平上比较保守，且叶绿体基因组可在居群水平上揭示物种的地理遗传结构。这为后续开展刺续断属物种群体遗传学和系统发育基因组学研究奠定了基础。

关键字：白花刺续断，叶绿体全基因组，药用植物，基因，系统发育基因组学

中图分类号：　　文献标识码：A　　文章编号：1000-3142（2021）00-0000-00

Analysis of complete chloroplast genomes from wild populations of *Acanthocalyx alba*

ZHANG Qian[1], ZHANG Dequan[1,2]*

(1. *College of Pharmacy*, *Dali University*, Dali 671000, Yunnan, China; 2. *Yunnan Key Laboratory of Screening and Research on Anti-pathogenic Plant Resources from Western Yunnan*, Dali 671000, Yunnan, China)

Abstract: *Acanthocalyx alba* is a common medicinal plant in Tibetan region, and there are few reports on its whole chloroplast genome now. In order to reveal basic characteristics of the whole chloroplast genome of *Acanthocalyx alba* and explore its phylogeographic structure, ten individuals from five wild populations were sequenced by Illumina sequencing platform. After assembly and annotation, ten complete chloroplast genome sequences were obtained. Moreover, genome characteristics and phylogenetic relationships were explored. The results were as follows: (1) The complete chloroplast genomes of *Acanthocalyx alba* were 155 335 - 156 266 bp. A total of 113 genes were annotated, including 72 protein-coding genes, 30 tRNA genes and four rRNA

基金项目：国家自然科学基金项目（32060091）；云南省中青年学术和技术带头人后备人才项目（202105AC160063）；大理大学中药资源与民族药创新团队项目（ZKLX2019318）；云南省李剑专家工作站（202005AF150013）[Supported by National Natural Science Foundation of China (32060091); Reserve Talents Project for Young and Middle-Aged Academic and Technical Leaders of Yunnan Province（202105AC160063）; Dali University Chinese Medicine Resources and Ethnic Medicine Innovation Team Project (ZKLX2019318); Li Jian Expert Workstation of Yunnan Province (202005AF150013)]。

作者简介：张倩（1996-），女，在读硕士研究生，研究方向为药用植物基因组学研究，(E-mail)1349362994@qq.com。

通信作者：张德全，教授，硕士生导师，主要从药用植物的分类和生药的分子鉴定研究，(E-mail)zhangdeq2008@126.com。

genes. The chloroplast genome was highly conserved in species in terms of size, structure, GC content and gene composition; (2) Genomic comparative analysis showed that the fragments with large variation were located in the single copy region, and there was no obvious expansion and contraction at the IR boundaries; (3) Analysis of population genetics revealed rich genetic diversity in the wild populations and significant geographic structure, and there were certain relationships between genetic distances and geographic distances. This study lays the foundation for studies on population genetics and phylogenomics of species in *Acanthocalyx*.

Keywords: *Acanthocalyx alba*, complete chloroplast genome, medicinal plant, gene, phylogenomics

白花刺续断（*Acanthocalyx alba* (Hand.-Mazz.) M.J.Cannon），又名白花刺参，隶属于川续断科（Dipsacaceae）刺续断属（*Acanthocalyx*）。该属在我国有 4 种 2 变种，主要分布在云南、四川、西藏等地（Hong et al., 2011）。白花刺续断以全草入药，为传统藏药，其藏药名为“江才嘎保”，始载于《四部医典》（国家中医药管理局中华本草编委会，2002）。它是国家卫生部藏药药品标准中收载的三种“刺参”之一（青海省药品检验所和青海省藏医药研究所，1996），具有健胃、催吐之功效。内服可用于关节疼痛、小便失禁、腰痛、眩晕及口眼歪斜，外用治疗疮、化脓性创伤，还具有抗肿瘤作用（国家中医药管理局中华本草编委会，2002；杨竞生，1989）。近年来，有关白花刺续断的研究主要集中在其有效成分、含量测定及提取工艺等方面（吴春蕾等，2011；Zhang et al., 2013；杨圣贤等，2014；张志锋等，2015）。如张志锋等（2018）发现白花刺续断中含有皂苷、生物碱、甾醇等类型化合物，其中皂苷类是其主要有效成分。而分子生物学方面的研究较少，仅有 Wang 等（2020）报道了该种的叶绿体基因组序列。那么，在白花刺续断的种内居群水平上，其叶绿体基因组序列有什么变化特征呢？

叶绿体基因组在被子植物中通常为母系遗传。与核基因组和线粒体基因组相比，其基因进化速率慢，且在基因组成及结构方面比较保守（Smith, 2015；Szymon et al., 2016；Du et al., 2020），这使得叶绿体基因组在植物物种鉴定和谱系进化研究具有重要作用。Cui 等（2019）对比分析了 32 种豆蔻属（*Amomum*）植物叶绿体基因组，结果表明叶绿体全基因组可准确鉴定豆蔻属物种；李依容等（2020）利用叶绿体基因组揭示了民族药滇白珠（*Gaultheria leucocarpa* var. *yunnanensis*）复合群的系统发育关系；Zhang 等（2021）基于叶绿体基因组重建了桃金娘目（Myrtiflorae）的物种分化时间和系统发育关系。但叶绿体基因组在长期的进化过程中，结构上会发生一些变异，如反向重复区收缩、倒位、基因和内含子的丢失等（Zhang et al., 2014；Liao et al., 2020；姜汶君等，2020），这些结构变异为揭示物种系统发育和进化关系等提供了遗传信息。由此可见，植物叶绿体全基因组序列能提供丰富的遗传信息，在分类鉴定、系统发育及进化等方面具有重要意义。那么，植物叶绿体全基因组是否适用于种内居群水平上的群体遗传学研究？由于群体水平上测序成本较高，数据分析方法尚不成熟等，这方面的相关研究还比较少。

本研究，以白花刺续断野生居群个体为研究材料，拟采用二代测序技术进行高通量测序，并对其叶绿体全基因组进行拼接、注释及进化分析。拟探讨以下科学问题：（1）白花刺续断的叶绿体全基因组序列有什么特征？（2）叶绿体全基因组能否用于解析白花刺续断种内居群水平上的遗传结构？通过该研究，将为刺续断属相关物种的分子遗传学研究奠定基础，也为叶绿体全基因组在群体遗传研究方面开展初步尝试。

1　材料与方法

1.1　实验材料

本实验中，白花刺续断分子材料采自于四川省甘孜州的 5 个野生居群（表 1），共 10 份实验材料。经大理大学张德全教授鉴定为白花刺续断（*Acanthocalyx alba* (Hand.-Mazz.) M.J.Cannon），其凭证标本保存于大理大学药学院药用植物与生药标本馆。

表 1 白花刺续断样品采集信息

Table 1 Collection information of *Acanthocalyx alba*

编号 Number	采集地点 Locality	凭证标本 Voucher specimen	经纬度 Latitude/Longitude	海拔 Altitude (m)
SD01	四川省稻城县桑堆 Sangdui, Daocheng, Sichuan	ZDQ17115	29°07′53.41″/100°10′21.41″	4 005
SD02	四川省稻城县桑堆 Sangdui, Daocheng, Sichuan	ZDQ17115	29°07′53.41″/100°10′21.41″	4 005
YD01	四川省稻城县亚丁机场 Yading Airport, Daocheng, Sichuan	ZDQ17105	29°15′17.87″/100°05′02.76″	3 823
YD02	四川省稻城县亚丁机场 Yading airport, Daocheng, Sichuan	ZDQ17105	29°15′17.87″/100°05′02.76″	3 823
DF01	四川省道孚县 Daofu, Sichuan	ZDQ17178	30°44′59.62″/101°19′51.19″	3 850
DF02	四川省道孚县 Daofu, Sichuan	ZDQ17178	30°44′59.62″/101°19′51.19″	3 850
LT01	四川省理塘县 Litang, Sichuan	ZDQ17124	29°29′04.73″/100°13′40.35″	4 510
LT02	四川省理塘县 Litang, Sichuan	ZDQ17124	29°29′04.73″/100°13′40.35″	4 510
KD01	四川省康定县 Kangding, Sichuan	ZDQ16035	30°02′51.00″/101°24′01.57″	3 884
KD02	四川省康定县 Kangding, Sichuan	ZDQ16035	30°02′51.00″/101°24′01.57″	3 884

1.2 基因组 DNA 提取与测序

DNA 提取采用改良的 CTAB 方法，从硅胶干燥的叶片材料中提取总基因组 DNA。利用 Covaris 超声波破碎仪将基因组 DNA 片段化，经末端修复、加 A 尾、加测序接头、纯化及 PCR 扩增等过程，构建测序文库。文库经质量检测合格后，使用带有双末端 pair-end (2×300bp) Illumina Hiseq 2 500 平台进行二代测序，整个测序工作在北京诺禾致源生物科技有限公司完成。

1.3 叶绿体全基因组的组装、注释

经二代测序，得到 4G 左右原始数据（Raw Data），经 Trimmomatic v.0.32 过滤处理后，利用 GetOrganelle.py 进行组装，后续数据处理参考本课题组前期工作（胡海粟和张德全，2021）。以白花刺续断 *Acanthocalyx alba*（序列号：NC_045055）为参考基因组，使用 Geneious 8.0.2 软件完成白花刺续断叶绿体全基因组的注释，并将其提交到 NCBI 数据库（https://www.ncbi.nlm.nih.gov/），利用在线工具 Organellar Genome Draw（https://chlorobox.mpimp-golm.mpg.de/OGDraw.html）绘制叶绿体基因组物理图谱。

1.4 IR/SC 边界收缩扩张及 SSR 分析

将注释好的 10 条基因组序列上传至网站 IRscope（https://irscope.shinyapps.io/irapp/）进行 IR 边界的收缩和扩张分析，最后得到的图片采用绘图工具 Adobe Illustrator CC 2015 进行人工调整。同时，利用 MISA 软件搜索简单重复序列（SSR: Simple sequence repeats）位点并分析白花刺续断叶绿体基因组的 SSR 特征。

1.5 序列差异比较分析

在采用 MAFFT v.7.129 软件对 10 条白花刺续断叶绿体基因组序列进行比对后，使用

BioEdit 软件手动调整序列。使用 DnaSP v.7.0.26 对叶绿体基因组中的核苷酸变异性（Pi）进行滑动窗口分析。步长设置为 200 bp，窗口长度为 600 bp。P-distance 使用 MEGA v.7.0.26 软件进行统计分析。此外，将叶绿体基因组序列的注释进行格式转换，利用在线软件 mVISTA（http://genome.lbl.gov/vista/mvista/submit.shtml）的 Shuffle-LAGAN 模式对白花刺续断叶绿体全基因组比较分析，选取 *Acanthocalyx alba*（NC_045055）作为参考序列。

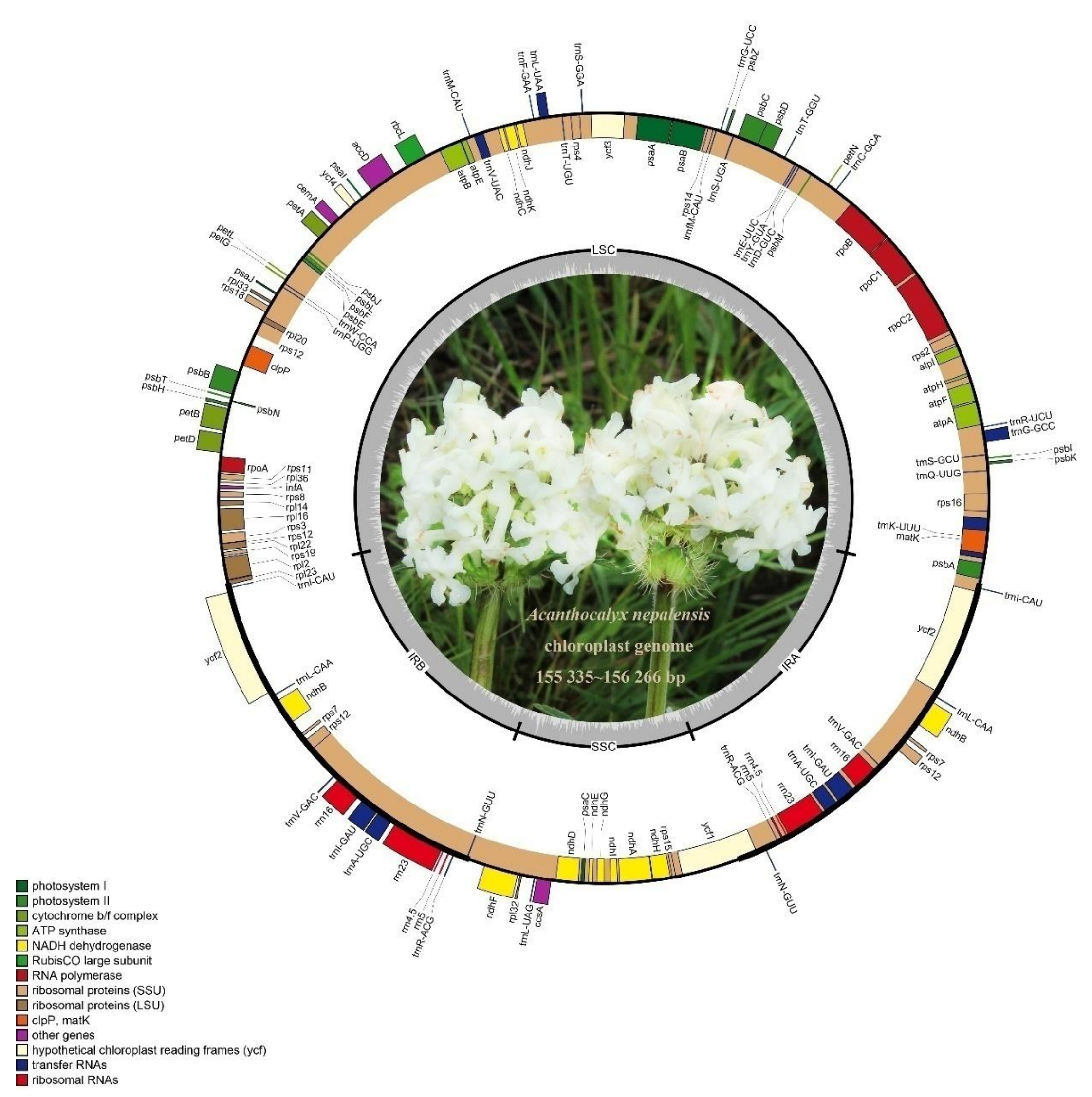
1.6 系统发育分析

根据白花刺续断叶绿体基因组注释信息，从 NCBI 数据库下载已发表的川续断科及忍冬科（Caprifoliaceae）的 11 种植物的叶绿体全基因组序列，用于系统发育分析。选择小粒咖啡（*Coffea arabica* L.）和中粒咖啡（*C. arabica* Pierre *ex* Froehn.）为外类群，使用 MAFFT V.7.1 将白花刺续断与下载的叶绿体全基因组序列进行多序列比对。系统发育树的构建，采用了最大似然法（Maximum likelihood, ML）、最大简约法（Maximum parsimony, MP）和贝叶斯推论法（Bayesian inference, BI）三种方法。核苷酸替代模型经 jModelTest v 2.1.7 软件筛选定为 GTR+G 模型。利用 RAxML v.8.2.4 软件构建 ML 系统树，采用快速靴带算法，重复 1 000 次。利用 MEGA v.7.0.26 软件构建 MP 树，重复 1 000 次。利用 MrBayes v.3.2.6 构建 BI 树，基于马尔科夫链蒙特卡洛（MCMC）算法，计算 100 万代，每隔 1 000 代取样一次，舍弃前 25%棵树，根据剩余的样本构建一致树。

2 结果与分析

2.1 基因组结构与基本特征

白花刺续断的叶绿体全基因组为常见的四分体结构，由两个反向重复区 IRs（Inverted repeats）、一个大单拷贝区 LSC（Large single copy）和一个小单拷贝区 SSC（Small single copy）组成（表 2，图 1）。拼接后的白花刺续断叶绿体基因组，全长为 155 335~156 266 bp，GC 含量为 38.1~38.2%。各区段长度分别为：89 027~89 076 bp（LSC）、17 689~17 842 bp（SSC）、24 253~24 666 bp（IRs）。4 个区段中 GC 含量最高的是 IR 区（42.8~43.2%），其次是 LSC 区（36.5%）和 SSC 区（32.9%）。经注释，得到 113 个基因，包括 72 个编码蛋白基因、30 个 tRNA 基因、4 个 rRNA 基因和 7 个假基因（*clp*P、*acc*D、*ycf*2、*ycf*1、*rps*18、*rps*3 和 *ycf*3）。此外，白花刺续断叶绿体全基因组中有 16 个基因含有内含子（intron），且均只含有一个内含子（表 3）。



圆圈外的基因是顺时针转录，圆圈内的基因是逆时针。图中颜色表示功能基因。内部的深灰色对应于 GC 含量，浅灰色对应于 AT 含量。

The genes outside the circle are transcribed clockwise, while the genes inside the circle are transcribed counterclockwise. The colors here represent functional genes. The inner dark gray corresponds to the GC content and the light gray corresponds to the AT content.

图 1 白花刺续断叶绿体全基因组图谱

Fig.1 Gene map of complete chloroplast genome in *Acanthocalyx alba*

表 2 白花刺续断 10 条叶绿体全基因组序列特征的比较

Table 2 Comparison on ten chloroplast complete genomes in *Acanthocalyx alba*

编号 Number	总长 Total (bp)	大单拷贝区 LSC (bp)	小单拷贝区 SSC (bp)	反向重复区 IR (bp)	GC%	基因数 Total genes	蛋白编码基因 CDS	tRNA	rRNA	GenBank 登录号 Accession number
SD01	156 082	89 030	17 740	24 656	38.1	113	72	30	4	MW053688
SD02	155 335	89 035	17 794	24 253	38.2	113	72	30	4	MW053689
YD01	156 110	89 036	17 842	24 616	38.1	113	72	30	4	MW053690
YD02	156 053	89 036	17 705	24 656	38.1	113	72	30	4	MW053687
DF01	156 266	89 027	17 713	24 763	38.1	113	72	30	4	MW053686
DF02	156 266	89 027	17 713	24 763	38.1	113	72	30	4	MW053685
LT01	156 054	89 032	17 790	24 616	38.1	113	72	30	4	MW018701
LT02	156 054	89 032	17 790	24 616	38.1	113	72	30	4	MW018702
KD01	156 097	89 076	17 689	24 666	38.1	113	72	30	4	MW018703
KD02	156 097	89 076	17 689	24 666	38.1	113	72	30	4	MW018704

表 3 白花刺续断叶绿体全基因组基因组成

Table 3 Composition of complete chloroplast genome of *Acanthocalyx alba*

基因类别 Gene group	基因功能 Gene function	基因名称 Gene name
自我复制 Self-replication	核糖体大亚基 Large subunit of ribosone	*rpl*2，*rpl12*^{I}*，*rpl*32，*rpl*22，*rpl*16*，*rpl*14，*rpl*36，*rpl*20，*rpl*33
	核糖体小亚基 Small subunit of ribosone	*rps*7^{I}，*rps*15，*rps*19，*rps*3$^{\#}$，*rps*8，*rps*11，*rps12*^{I}*，*rps*18$^{\#}$，*rps*4，*rps*14，*rps*2，*rps*16*
	RNA 聚合酶 RNA polymerase	*rpo*A，*rpo*B，*rpo*C1*，*rpo*C2
	rRNA	*rrn16*^{I}，*rrn23*^{I}，*rrn4.5*^{I}，*rrn5*^{I}
	tRNA	*trn*I-CAU^{I}，*trn*L-CAA^{I}，*trn*V-GAC^{I}，*trn*I-GAU^{I}*，*trn*A-UGC^{I}*，*trn*R-ACG^{I}，*trn*N-GUU^{I}，*trn*L-UAG，*trn*P-UGG，*trn*W-CCA，*trn*M-CAU，*trn*V-UAC*，*trn*F-GAA，*trn*L-UAA*，*trn*T-UGU，*trn*S-GGA，*trnf*M-CAU，*trn*G-UCC，*trn*S-UGA，*trn*T-GGU，*trn*E-UUC，*trn*Y-GUA，*trn*D-GUC，*trn*C-GCA，*trn*R-UCU，*trn*G-GCC*，*trn*S-GCU，*trn*Q-UUG，*trn*K-UUU*，*trn*H-GUG
光合作用 Photosynthesis	光学系统 1 Photosystem I	*psa*C，*psa*J，*psa*I，*psa*A，*psa*B
	光学系统 2 Photosystem II	*psb*H，*psb*N，*psb*T，*psb*B，*psb*E，*psb*F，*psb*L，*psb*J，*psb*Z，*psb*C，*psb*D，*psb*M，*psb*I，*psb*K，*psb*A
	NADH-脱氢酶 NADH dehydrogenase	*ndh*B^{I}*，*ndh*H，*ndh*A*，*ndh*I，*ndh*G，*ndh*E，*ndh*D，*ndh*F，*ndh*C，*ndh*K，*ndh*J
	细胞色素复合物 Cytochrome b/f complex	*pet*D*，*pet*B*，*pet*G，*pet*L，*pet*A，*pet*N
	ATP 合酶 ATP synthase	*atp*B，*atp*E，*atp*I，*atp*H，*atp*F*，*atp*A
	二磷酸核酮糖羧化酶大亚基基因 Large subunit of rubisco	*rbc*L
其他基因 Other genes	翻译起始因子 Translational initation factor	*inf*A
	成熟酶基因 Maturase	*mat*K
	蛋白酶基因 Protease	*clp*P*$^{\#}$
	包膜蛋白基因 Envelope membrane protein	*cem*A
	乙酰 CoA 羧化酶的亚基 Subunit of Acetyl-carboxylase	*acc*D$^{\#}$
	C 型细胞色素合成 C-type cytochrome synthesis	*ccs*A
	开放阅读框 Open reading frames	*ycf*2$^{I\#}$，*ycf*1$^{\#}$，*ycf*4，*ycf*3$^{\#}$

注: “I”表示该基因位于 IR 区，“*”表示该基因含有内含子，“#”代表假基因。

Note: "I" indicates that the gene is located in the IR region, "*" indicates that the gene contains introns, and "#" represents a pseudogene.

2.2 IR/SC 边界收缩扩张及 SSR 分析

叶绿体基因组由两个反向重复的 IR 区、LSC 区与 SSC 区构成，因此存在 LSC/IRb、IRb/SSC、SSC/IRa 与 IRa/LSC 四个边界。在基因组进化过程中，4 个边界会发生扩张与收缩，使某些基因进入 IR 区或单拷贝区。不同地点白花刺续断的叶绿体基因组的 4 个边界相对保守（图 2）。LSC/IRb 边界在白花刺续断中位于 *rpl*23 基因内部，且位于 LSC 区域的差异不大，分别是 185~186 bp；SD01、YD01、YD02、LT01、LT02、KD01、KD02 的 IRb/SSC 边界基因完全相同，位于 IRb 区 *trn*N-GUU 基因 138 bp 处，而 SD02、DF01、DF02 的 *trn*N-GUU 基因扩张到 SSC 内部，距离 IRb/SSC 边界 48~223 bp；SSC/IRa 的边界在白花刺续断基因组中都位于 *ycf1* 基因内部；IRa/LSC 边界全部位于 *trn*H-GUG 基因附近。

利用 MISA 软件对白花刺续断叶绿体全基因组进行分析，在白花刺续断 10 条序列中分别检测到 70、68、70、70、74、74、70、70、71、71 个 SSR 位点（图 3B）。SSR 最丰富的类型为单核苷酸重复，其次是二核苷酸重复、三核苷酸重复、五核苷酸重复、四核苷酸重复和六核苷酸重复。白花刺续断叶绿体基因组中的 SSR 主要是由 A 和 T 组成，其中大部分是以 A/T 碱基构成的单核苷酸重复，其次是由 AT/TA 构成的二核苷酸重复（图 3A）。进一步分析表明，大部分 SSR 位于 LSC 区域，小部分位于 SSC 和 IR 区域（图 3C）。同时，基因组中的 SSR 大部分分布于基因间区（Intergenic spacer, IGS）中，其他少数 SSR 分布在内含子和蛋白编码基因（CDS）中（图 3D）。

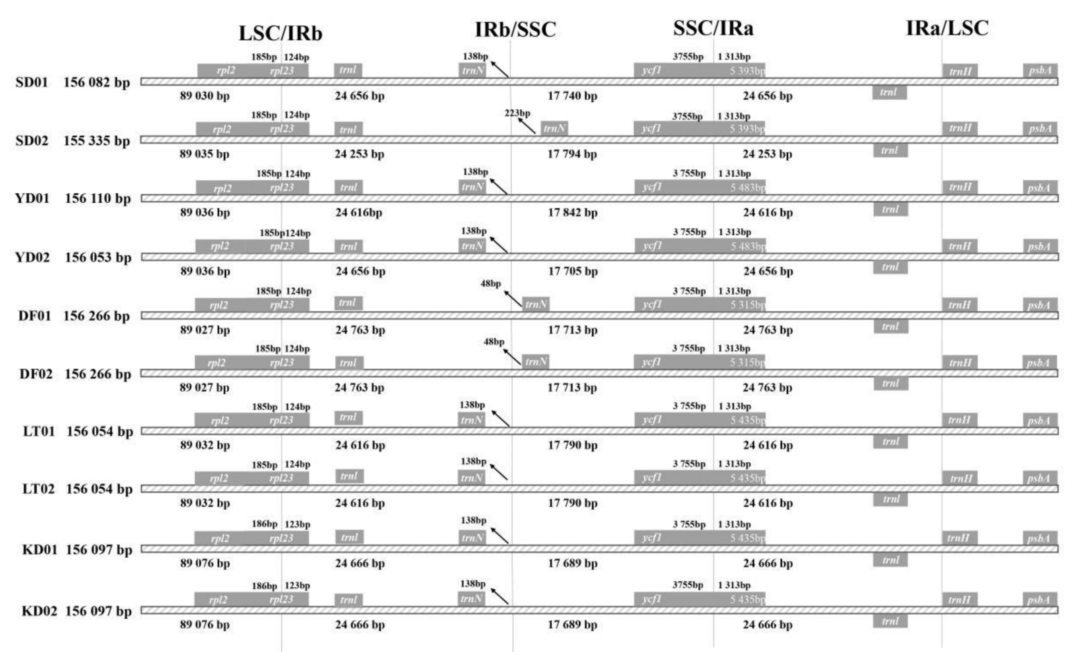
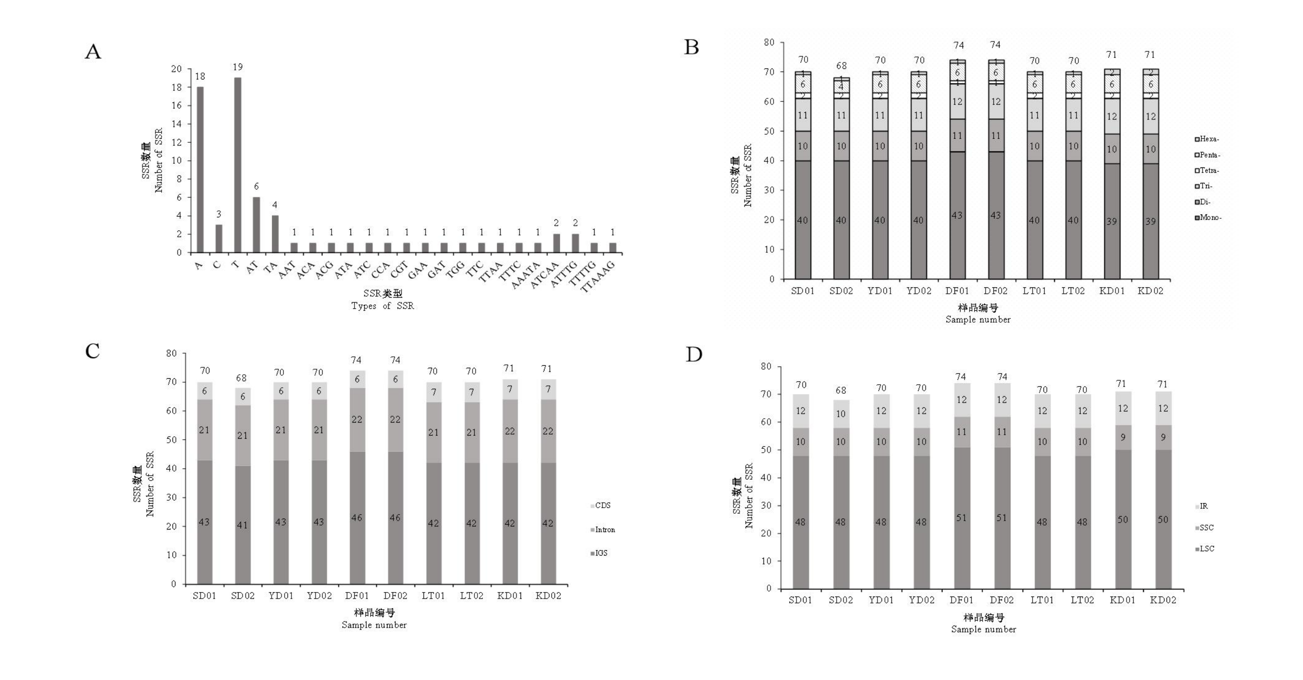


图 2 白花刺续断叶绿体基因组 LSC、SSC 和 IR 边缘区的比较

Fig. 2 Comparison of LSC, SSC, and IR border regions among ten chloroplast genomes in *Acanthocalyx alba*



A. SD01 叶绿体基因组中不同重复类型 SSR 出现频率；**B**. 10 个基因组中不同 SSR 类型的数目；**C**. 在 LSC、SSC 和 IR 区域中出现 SSR 的频率；**D**. 在 IGS、CDS 和内含子中出现 SSR 的频率。

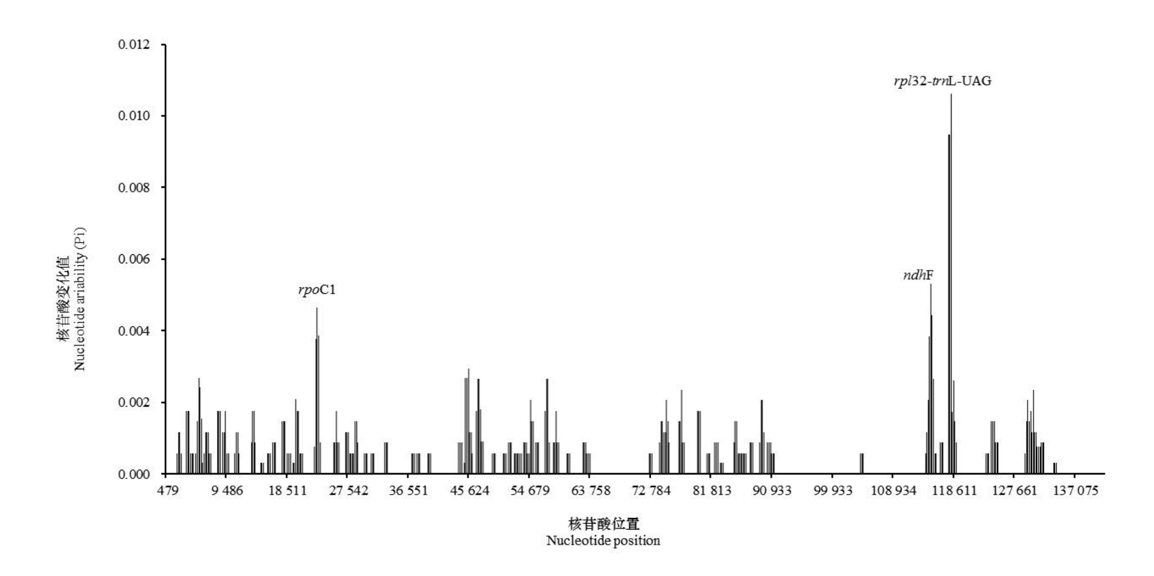
A. Frequency of SSR motifs in different repeat types of SD01 chloroplast genome; **B**. number different SSRs type detected in ten genomes; **C**. frequency of identified SSR in LSC, SSC, and IR regions; **D**. frequency of identified SSR in IGS, CDS, and Intron.

图 3 白花刺续断 10 条叶绿体基因组的 SSR 分析

Fig. 3 Analysis of simple sequence repeat (SSR) on ten chloroplast genomes in *Acanthocalyx alba*

2.3 序列差异比较分析

将比对好的 5 个地区白花刺续断叶绿体基因组进行 Sliding window 分析（图 4）。结果显示，SSC 区域的变异水平最高，IR 区域最低。同时，筛选到 3 条高变异序列，分别位于 LSC 区（*rpo*C1）和 SSC 区（*ndh*F 和 *rpl*32-*trn*L-UAG）。其中 *rpl*32-*trn*L-UAG 的变异性最高，其次是 *ndh*F，而 *rpo*C1 最低。此外，本研究中以 SD01 作为参考序列，与其余 9 条白花刺续断叶绿体基因组进行两两比较分析。结果显示，叶绿体基因组序列中非编码区变异高于蛋白编码区域，单拷贝区（LSC&SSC）变异明显大于反向重复区（IR）。5 个地区白花刺续断叶绿体基因组序列整体上高度相似，变异较大的基因有 *rpo*C2、*psb*C、*rrn*23 和 *ycf*1，其他基因保守程度非常高。基因间区的变异大于基因区，如 *atp*F-*atp*H、*psa*B-*psa*A、*psa*A-*ycf*3、*trn*M-CAU-*atp*E、*psb*F-*psb*E、*psb*E-*pet*L、*rrn*5-*trn*N-ACG、*trn*R-ACG-*trn*N-GUU、*trn*L-UAG-*ccs*A（图 5）。从这些区域中，可开发特异性片段，用于该属种间及种下水平的系统进化与发育研究。

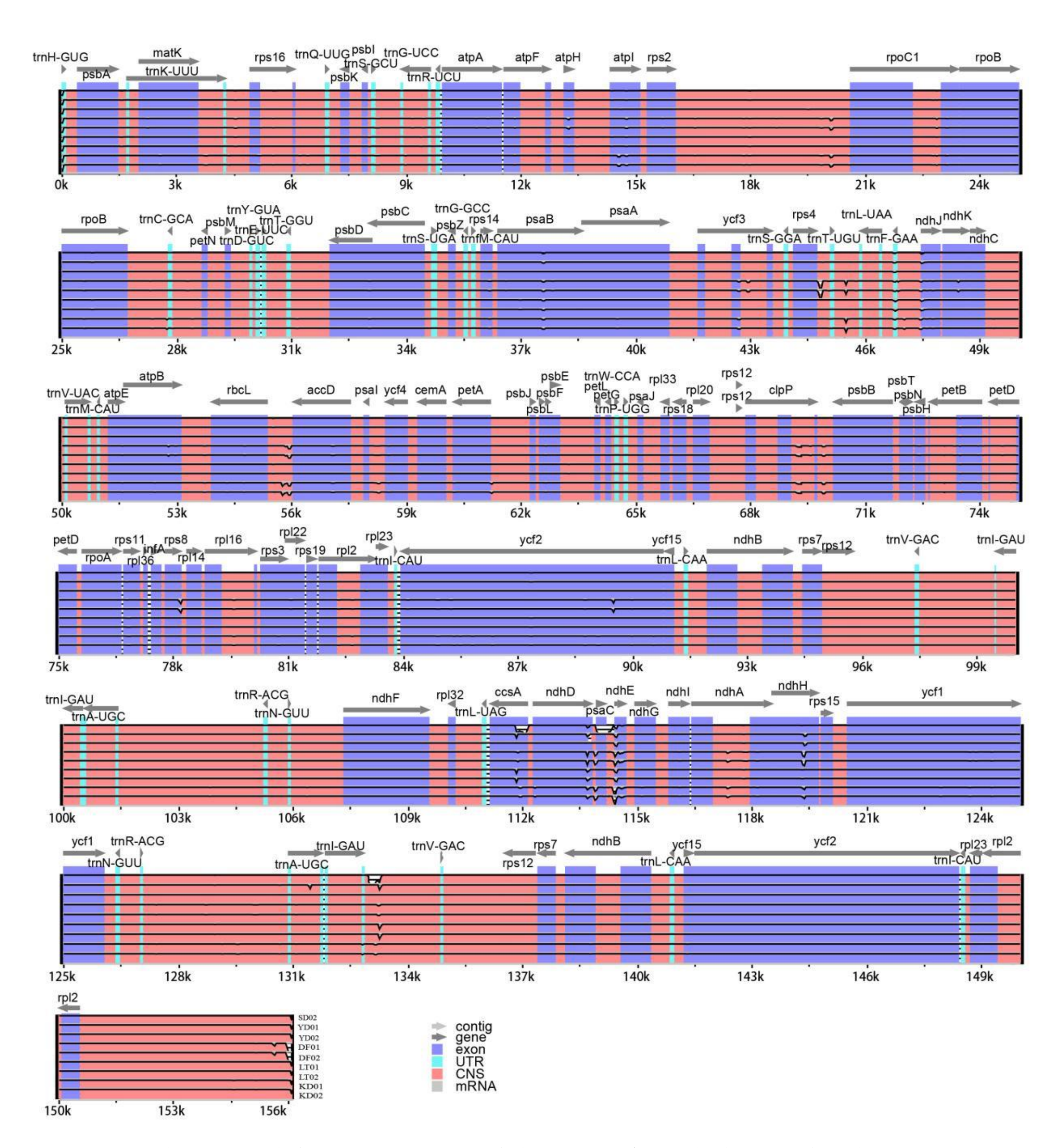


窗口长度：600bp，步长：200bp。X 轴：窗口中点的位置；Y 轴：每个窗口的核苷酸多样性。

Window length: 600 bp, step size: 200 bp. X-axis: position of the midpoint of a window; Y-axis: nucleotide diversity of each window.

图 4 白花刺续断 10 条叶绿体全基因组的滑动窗口分析

Fig. 4 Sliding window analysis of ten chloroplast genomes in *Acanthocalyx alba*



基于 VISTA 的标识图显示以 SD01 为参照，对九个白花刺续断进行序列鉴定。

VISTA-based identify plot showing sequence identify among nine *Acanthocalyx alba* using SD01 as a reference.

图 5 白花刺续断 10 条叶绿体基因组的可视化比对

Fig. 5 Visualization alignment of ten chloroplast genomes in *Acanthocalyx alba*

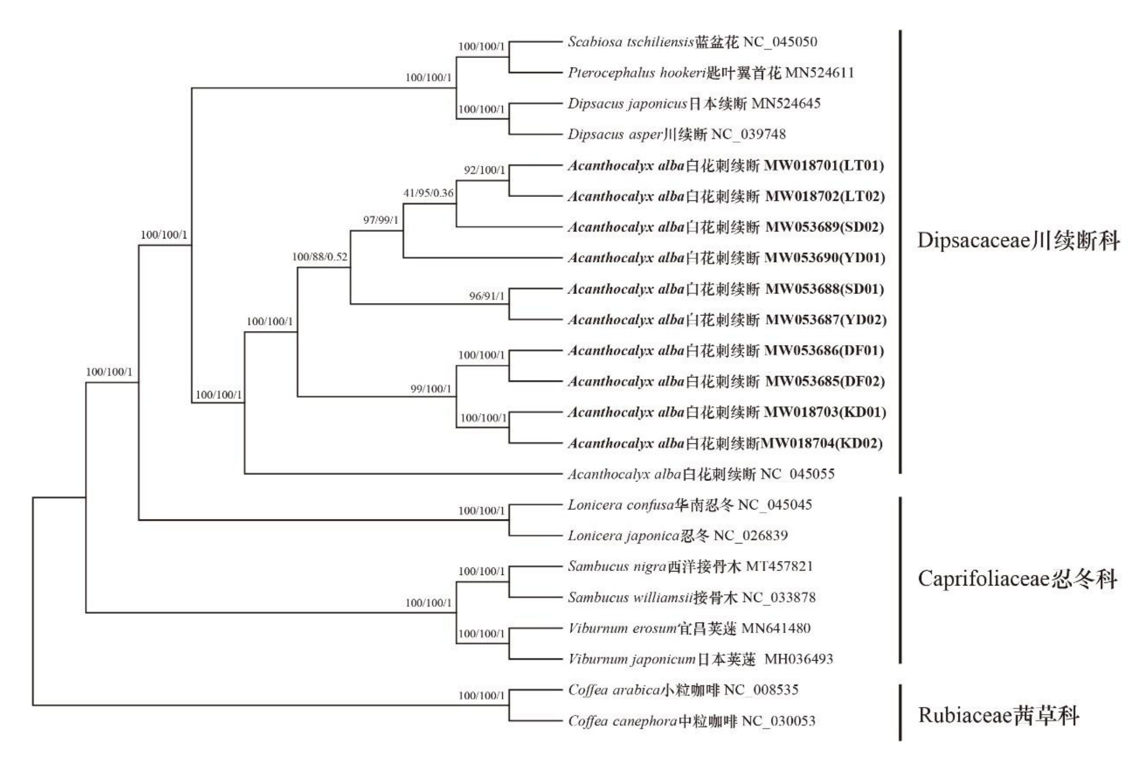
2.4　系统发育分析

本研究利用 P-distance 种间遗传变异及核苷酸替换比较了 10 条白花刺续断的叶绿体全基因组进化差异，研究结果表明，P-distance 为 0.000 0~0.000 7，核苷酸差异值为 0~1 515，且大部分序列间地理位置越远，其相互间 P-distance 和核苷酸差异值越大（表 4）。系统发育分析结果显示，三种方法构建的进化树所反映的不同野生居群之间的进化关系相似（图 6），也与遗传距离分析的结果相吻合。在系统发育树中，康定（KD）和道孚（DF）的四个个体最早分化出来，其次是亚丁（YD）和桑堆（SD），最后是理塘（LT）的两个个体。但亚丁（YD）和桑堆（SD）的四个个体不能明显分开。

表 4 白花刺续断个体间遗传距离与核苷酸差异值

Table 4 Genetic distances and nucleotide difference values among individuals of *Acanthocalyx alba*

	SD01	SD02	YD01	YD02	DF01	DF02	LT01	LT02	KD01	KD02
SD01		940	263	45	819	818	218	218	588	588
SD02	0.000 1		791	933	1 515	1 514	746	746	1 306	1 306
YD01	0.000 1	0.000 0		256	851	850	67	67	625	625
YD02	0.000 0	0.000 1	0.000 1		781	780	211	211	551	551
DF01	0.000 6	0.000 6	0.000 6	0.000 6		1	805	805	576	576
DF02	0.000 6	0.000 6	0.000 6	0.000 6	0.000 0		804	804	575	575
LT01	0.000 1	0.000 1	0.000 1	0.000 1	0.000 6	0.000 6		0	581	581
LT02	0.000 1	0.000 1	0.000 1	0.000 1	0.000 6	0.000 6	0.000 0		581	581
KD01	0.000 7	0.000 7	0.000 7	0.000 7	0.000 5	0.000 5	0.000 7	0.000 7		0
KD02	0.000 7	0.000 7	0.000 7	0.000 7	0.000 5	0.000 5	0.000 7	0.000 7	0.000 0	



上面的节点数是支持值，左边是 MP 自展值，中间是 ML 自展值，右边是贝叶斯后验概率（PP）值。图中加黑部分为本文主要研究对象。

Number above nodes are support values with MP bootstrap values on the left, ML bootstrap values in the middle, Bayesian posterior probabilities (PP) values on the right. The black part in the figure is the main research object of this paper.

图 6 利用最大简约法（MP）、最大似然法（ML）和贝叶斯分析法（BI）研究了 10 个白花刺续断叶绿体全基因组的系统发育关系

Fig. 6 Phylogenetic relationship of ten *Acanthocalyx alba* based on complete chloroplast genome using maximum parsimony (MP), maximum likelihood (ML), and Bayesian analyses (BI) ethods

3 讨论与结论

本研究报导了白花刺续断的叶绿体全基因组序列特征，并在居群水平上揭示了其地理遗传结构。不同野生居群的叶绿体基因组所编码的基因类别、数量及排列顺序高度一致。同时个体间具有高度相似的 GC 含量，单个序列中 IRs 区序列的 GC 含量最高。白花刺续断叶绿体基因组中共含有 7 个假基因，其中 5 个假基因是川续断科植物所共有的（*clp*P、*acc*D、*ycf*2、*ycf*1、*rps*18），故推测川续断科植物假基因现象可能是普遍存在的（Wang et al., 2020）。叶绿体 SSR 位点是一种高效的分子标记。本研究中，白花刺续断叶绿体全基因组序列的 SSRs 主要以 A/T 碱基为主，这与其它被子植物中的情况相似（Guo et al., 2017；Na et al., 2018；Chen et al., 2019）。同时，这也进一步证实了叶绿体 SSRs 主要是由 polyA 和 polyT 重复所构成，而较少含有 C 或 G 串联重复的观点（Kuang et al., 2011）。此外，这些 SSRs 主要分布在两个单拷贝区，故推测这些高 A/T 含量的 SSRs 和分布于 IR 区的 rRNA 序列可能是导致叶绿体基因组中 GC 含量偏低以及各区域碱基含量差异的潜在原因（张明英等，2020）。

IR区和SC区的扩张和收缩被认为是直接影响被子植物叶绿体基因组大小的重要因素（Wang et al., 2017；Song et al., 2019）。本研究表明，不同野生居群个体间叶绿体基因组4个边界均未出现明显的扩张和收缩现象，说明白花刺续断叶绿体基因组IRs区大小高度保守，

这也与Wang等（2020）的研究结果一致。从叶绿体基因组中发掘的高变片段，不仅可以在物种水平上用于系统发育和物种鉴定研究，也可以在居群水平上提供丰富的遗传信息，从而揭示物种的居群动态与进化历史等。Fatemeh等（2018）基于*rpl*32-*trn*L-UAG对滇紫草属（*Onosma*）物种进行系统发育分析和分化时间估计；Nahla等（2020）采用*rpo*C1对苜蓿属（*Medicago*）植物进行亲缘关系分析；Chen等（2020）基于叶绿体基因组对贝母属（*Fritillaria*）植物进行研究，结果发现筛选出的*ycf*1和*psb*M-*psb*D可作为特定条形码用于贝母属植物物种鉴定。这些研究工作也进一步证实高变片段在物种进化及鉴定等方面具有特殊作用。本研究中，筛选出3个高变片段（*rpo*C1、*ndh*F和*rpl*32-*trn*L-UAG），可用于刺续断属内种间系统发育及种内群体遗传学研究。

传统上，常用叶绿体基因片段来研究物种的群体遗传结构和谱系进化关系，但因叶绿体片段多态位点不足而作用有限（Zhang et al., 2019；Zhang et al., 2020；刘家奇等，2021）。与之相比，叶绿体全基因组具有极为丰富的遗传变异，为复杂植物类群的遗传进化研究提供有效手段。Wang 等（2020）基于美国山核桃（*Carya illinoinensis*）两个不同居群间叶绿体基因组核苷酸差异性，揭示了该物种居群水平的遗传多样性。本研究中，白花刺续断 5 个野生居群间具有较为明显的遗传结构，个体间的遗传距离、核苷酸差异值与地理距离之间呈较好的相关性。这也与系统发育树所揭示的进化关系相吻合。值得注意的是，桑堆（SD）和亚丁（YD）的四个个体没有形成独立分支，这可能是两个居群间地理距离较近引起相对频繁的基因流所致。这一结果也说明，与核基因组相比，叶绿体基因组进化较慢及单亲遗传的特性，其作用也有明显的局限性。

综上所述，叶绿体全基因组序列具有极为丰富的遗传信息，可为复杂植物类群及种下居群水平上的群体遗传及谱系进化研究提供有效手段。但由于二代测序的价格仍旧较为高昂，本文中居群样本量较少，该研究结果的科学性尚有不足。因此，叶绿体基因组能否作为传统的分子片段或标记的技术补充，需要更多研究工作来验证。此外，将叶绿体全基因组用于群体遗传学分析的数据分析方法也有待于进一步完善。

参考文献

CHEN Q, WU XB, ZHANG DQ, 2019. Phylogenetic analysis of *Fritillaria cirrhosa* D. Don and its closely related species based on complete chloroplast genomes[J]. Peer J, 7: e7480.

CHEN Q, WU XB, ZHANG DQ, 2020. Comparison of the abilities of universal, super, and specific DNA barcodes to discriminate among the original species of Fritillariae cirrhosae bulbus and its adulterants[J]. PLoS ONE, 15(2): e0229181.

CUI Y, CHEN X, NIE L, et al., 2019. Comparison and phylogenetic analysis of chloroplast genomes of three medicinal and edible *Amomum* species[J]. Int J Mol Sci, 20(16): 4040.

DU X, ZENG T, FENG Q, et al., 2020. The complete chloroplast genome sequence of yellow mustard (*Sinapis alba* L.) and its phylogenetic relationship to other Brassicaceae species[J]. Gene, 731: 144340.

FATEMEH N, SHAHROKH KO, VALEYOLLAH M, et al., 2018. Molecular phylogeny and divergence times of *Onosma* (Boraginaceaes.s.) based on nrDNA ITS and plastid *rpl*32-*trn*L (UAG) and *trn*H–*psb*A sequences[J]. Nord J Bot, 37(1): e02060.

GUO HJ, LIU JS, LUO L, et al., 2017. Complete chloroplast genome sequences of *Schisandra chinensis*: genome structure, comparative analysis, and phylogenetic relationship of basal angiosperms[J]. Sci Chin, 60(11): 1286-1290.

HONG DY, BARRIE FR, 2011. *Acanthocalyx* [M] // Dipsacaceae. Flora of China: Vol. 19. Beijing:

Science Press; St. Louis: Missouri Botanical Garden: 649-650.

HU HS, ZHANG DQ, 2021. DNA super-barcoding of several medicinal species in *Gentiana* from Yunnan province[J]. Chin J Chin Mater Med, 46（20）: 5260-5269. [胡海粟，张德全，2021. 几种滇产龙胆属 *Gentiana* 药用植物的 DNA 超级条形码研究[J]. 中国中药，46（20）: 5260-5269.]

JIANG WJ, GUO MY, PANG XH, 2020. Application of chloroplast genome in identification and phylogenetic analysis of medicinal plants[J]. World Chin Med, 15(5): 702-708,716. [姜汶君，郭梦月，庞晓慧，2020. 叶绿体基因组在药用植物鉴定及系统进化研究中的应用[J]. 世界中医药，15(5): 702-708,716.]

KUANG DY, WU H, WANG Y L, et al., 2011. Complete chloroplast genome sequence of *Magnolia kwangsiensis* (Magnoliaceae): implication for DNA barcoding and population genetics[J]. Genome, 54(8): 663-673.

LI YR, TANG LP, YANG SD, et al., 2020. Phylogeny of an ethnical medicinal plant, *Gaultheria leucocarpa* var. *yunnanensis* complex based on chloroplast DNA data[J]. Guihaia, 40(1): 71-82. [李依容，唐丽萍，杨淑达，等，2020. 基于叶绿体基因证据的民族药滇白珠复合群系统发育关系[J]. 广西植物，40(1): 71-82.]

LIAO YY, LIU Y, LIU X, et al., 2020. The complete chloroplast genome of *Myriophyllum spicatum* reveals a 4-kb inversion and new insights regarding plastome evolution in Haloragaceae[J]. Ecol Evol, 10(6): 1-13.

LIU JQ, FU MJ, WU HY, et al., 2021. Phylogeography of the loess plateau endemic plant *Prinsepia uniflora* inferred from the chloroplast DNA[J]. Guihaia, 41(3): 396-403. [刘家奇，付孟娇，吴海洋，等，2021. 基于叶绿体 DNA 的黄土高原特有植物蕤核的谱系地理学[J]. 广西植物，41(3): 396-403.]

NA T, HAN LM, CHEN C, et al., 2018. The complete chloroplast genome sequence of *Epipremnum aureum* and its comparative analysis among eight Araceae species[J]. PloS One, 13(3): e0192956.

NAHLA ES, MOHAMED I, 2020. Implications of *rbc*L and *rpo*C1 DNA barcoding in phylogenetic relationships of some Egyptian *Medicago sativa* L. Cultivars[J]. Egypt J Bot, 60(2): 451-460.

Qinghai Institute for drug control, Qinghai Institute of Tibetan medicine, 1996. Chinese Tibetan medicine[M]. Shanghai: Shanghai Science and Technology Press: 170-173. [青海省药品检验所，青海省藏医药研究所，1996. 中国藏药，上海：上海科学技术出版社，170-173.]

SMITH, 2015. Mutation rates in plastid genomes: they are lower than you might think[J]. Genome Biol. Evol. 7(5): 1227-1234.

SONG X, TING S, WEN JL, et al., 2019. Comparative analysis of the complete chloroplast genome among *Prunus mume*, *P. armeniaca*, and *P. salicina*[J]. Hortic Res, 6(22): 2015-2035.

State administration of traditional Chinese medicine Chinese herbal editorial board, 2002. Chinese Herb. Tibetan Medicine Volume[M]. Shanghai: Shanghai Science and Technology Press: 208-209. [国家中医药管理局中华本草编委会，2002. 中华本草.藏药卷[M]. 上海：上海科学技术出版社: 208-209.]

SZYMON AO, EWELINA Ł, TOMASZ K, et al., 2016. Chloroplasts: state of research and practical applications of plastome sequencing[J]. Planta, 244(3): 517-527.

WANG HX, LIU H, MOORE M, et al., 2020. Plastid phylogenomic insights into the evolution of

the Caprifoliaceae s.l. (Dipsacales)[J]. Mol Phylogenet Evol, 142: 106641.

WANG WB, YU H, WANG JH, et al., 2017. The complete chloroplast genome sequences of the medicinal plant *Forsythia suspensa* (Oleaceae)[J]. Int J Mol Sci, 18(11): 2288-2304.

WANG XW, RHEIN HS, JENKINS J, et al., 2020. Chloroplast genome sequences of *Carya illinoinensis* from two distinct geographic populations[J]. Tree Genet Genomes, 16(5): 62-75.

WU CL, LIU Y, ZHANG ZF, et al., 2011. Optimization of extracting condition for *Morina nepalensis* by uniform design[J]. Chin Trad Herb Drug, 42(12): 2447-2450. [吴春蕾，刘圆，张志锋，等，2011. 均匀设计法优选白花刺参的提取工艺研究[J]. 中草药，42(12): 2447-2450.]

YANG JS, 1989. Diqing Tibetan medicine[M]. Kunming: Yunnan Nationalities Publishing House: 409-410. [杨竞生，1989. 迪庆藏药，昆明：云南民族出版社: 409-410.]

YANG SX, WANG JX, ZHANG ZF, 2014. Microscopic study of traditional Tibetan medicine *M. nepalensis* D. Don var. *alba* (Hand-Mazz.) Y.C. Tang[J]. J Southwest Univ Nationalities ·Nat Sci Ed, 40(5): 706-710. [杨圣贤，王静霞，张志锋，2014. 藏药材江才嘎保的纤维特征研究[J]. 西南民族大学学报 • 自然科学版，40(5): 706-710.]

ZHANG MY, WANG GZ, ZHOU JH, et al., 2020. Development and characterization of 20 novel EST-SSR markers for *Pteroceltis tatarinowii*, a relict tree in China[J]. Appl Plant Sci, 8(2): e11320.

ZHANG MY, ZHANG YQ, LI YM, et al., 2020. Complete plastid genomes of *Bupleurum chinense* DC. and *B. boissieuanum* H.Wolff, with comparative and phylogenetic analyses of medicinal *Bupleurum* species[J]. Acta Pharm Sin, 56(2): 618-629. [张明英，张雨曲，李依民，等，2020. 北柴胡、紫花阔叶柴胡叶绿体全基因组解析及柴胡属药用植物叶绿体基因组比较与系统发育分析[J]. 药学学报，56(2): 618-629.]

ZHANG X, SU HL, YANG J, et al., 2019. Population genetic structure, migration, and polyploidy origin of a medicinal species *Gynostemma pentaphyllum* (Cucurbitaceae)[J]. Ecol Evol, 9(19): 1-26.

ZHANG Y, LI L, YAN TL, et al., 2014. Complete chloroplast genome sequences of *Praxelis* (*Eupatorium catarium* Veldkamp), an important invasive species[J]. Gene, 549: 58-69.

ZHANG XF, LANDIS JB, WANG HX, et al., 2021. Comparative analysis of chloroplast genome structure and molecular dating in Myrtales[J]. BMC Plant Biol, 21(1): 219.

ZHANG ZF, LV LY, LUO P, et al., 2013. Two new ursolic acid saponins from *Morina nepalensis* var. *alba* Hand-Mazz.[J]. Nat Prod Res, 27(24): 2256-2262.

ZHANG ZF, LV LY, LIU Y, et al., 2015. Simultaneous determination of two new triterpenoid saponins in *Morina nepalensis* by HPLC[J]. West Chin J Pharm Sci, 30(6): 724-726. [张志锋，吕露阳，刘圆，等，2015. HPLC 同时测定白花刺参中的两个新三萜皂苷[J]. 华西药学杂志，30(6): 724-726.]

ZHANG ZF, WU YJ, CHEN L, et al., 2018. Anti-inflammatory effects and mechanism of total saponins of Tibetan medicine *Morina nepalensis* var. *alba*[J]. Pharmacolo Clin Chin Mat Med, 34(5): 47-50. [张志锋，吴尤娇，陈丽，等，2018. 藏药材白花刺参总皂苷的抗炎作用及机制[J]. 中药药理与临床，34(5): 47-50.]